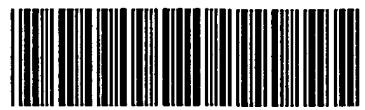# RAW SEQUENCE LISTING

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) no errors detected.

Application Serial Number: _09/943,857A_
Source: _IFW16_
Date Processed by STIC: _2-25-05_

ENTERED

IFW16

RAW SEQUENCE LISTING   DATE: 02/25/2005
PATENT APPLICATION: US/09/943,857A   TIME: 14:43:30

Input Set : A:\08919-066001.TXT
Output Set: N:\CRF4\02252005\I943857A.raw

```
 4 <110> APPLICANT: Shaw, Jei-Fu
 5         Lee, Guan-Chiun
 6         Tang, Shye-Jye
 9 <120> TITLE OF INVENTION: RECOMBINANT CANDIDA RUGOSA LIPASES
12 <130> FILE REFERENCE: 08919-066001
14 <140> CURRENT APPLICATION NUMBER: US 09/943,857A
15 <141> CURRENT FILING DATE: 2001-08-31
17 <160> NUMBER OF SEQ ID NOS: 11
19 <170> SOFTWARE: FastSEQ for Windows Version 4.0
21 <210> SEQ ID NO: 1
22 <211> LENGTH: 1641
23 <212> TYPE: DNA
24 <213> ORGANISM: Candida rugosa
26 <220> FEATURE:
27 <221> NAME/KEY: CDS
28 <222> LOCATION: (1)...(1641)
30 <400> SEQUENCE: 1
31 tcg atg aat tca cgt ggc cca gcc ggc cgt ctc gga tcg gta ccc acc    48
32 Ser Met Asn Ser Arg Gly Pro Ala Gly Arg Leu Gly Ser Val Pro Thr
33  1               5                  10                  15
35 gcc acg ctc gcc aac ggc gac acc atc acc ggt ctc aac gcc att gtc    96
36 Ala Thr Leu Ala Asn Gly Asp Thr Ile Thr Gly Leu Asn Ala Ile Val
37              20                  25                  30
39 aac gaa aag ttt ctc ggc ata ccg ttt gcc gag ccg ccc gtg ggc agc   144
40 Asn Glu Lys Phe Leu Gly Ile Pro Phe Ala Glu Pro Pro Val Gly Ser
41          35                  40                  45
43 ctc cgc ttc aag ccg ccc gtg ccg tac tcg gcg tcg ctc aac ggc cag   192
44 Leu Arg Phe Lys Pro Pro Val Pro Tyr Ser Ala Ser Leu Asn Gly Gln
45      50                  55                  60
47 cag ttt acc tct tac ggc ccg tct tgc atg cag atg aac cct atg ggc   240
48 Gln Phe Thr Ser Tyr Gly Pro Ser Cys Met Gln Met Asn Pro Met Gly
49  65                  70                  75                  80
51 tcg ttt gag gac aca ctt ccc aag aat gcg ctt gac ttg gtg ctc cag   288
52 Ser Phe Glu Asp Thr Leu Pro Lys Asn Ala Leu Asp Leu Val Leu Gln
53                  85                  90                  95
55 tcc aag atc ttc caa gtg gtg ctt ccc aac gac gag gac tgt ctc acc   336
56 Ser Lys Ile Phe Gln Val Val Leu Pro Asn Asp Glu Asp Cys Leu Thr
57              100                 105                 110
59 atc aac gtg atc cgg ccg ccc ggc acc agg gcc agt gct ggt ctc ccg   384
60 Ile Asn Val Ile Arg Pro Pro Gly Thr Arg Ala Ser Ala Gly Leu Pro
61              115                 120                 125
63 gtg atg ctc tgg atc ttt ggc ggt ggg ttt gag ctt ggc ggc tcc agc   432
64 Val Met Leu Trp Ile Phe Gly Gly Gly Phe Glu Leu Gly Gly Ser Ser
```

```
65      130                    135                   140
67 ctc ttt cca gga gac cag atg gtg gcc aag agc gtg ctc atg ggt aaa     480
68 Leu Phe Pro Gly Asp Gln Met Val Ala Lys Ser Val Leu Met Gly Lys
69 145                    150                   155                   160
71 ccg gtg atc cac gtg agc atg aac tac cgc gtg gcg tca tgg ggg ttc     528
72 Pro Val Ile His Val Ser Met Asn Tyr Arg Val Ala Ser Trp Gly Phe
73                    165                   170                   175
75 ttg gcc ggc ccc gac atc cag aac gaa ggc agc ggg aac gcc ggc ttg     576
76 Leu Ala Gly Pro Asp Ile Gln Asn Glu Gly Ser Gly Asn Ala Gly Leu
77                    180                   185                   190
79 cat gac cag cgc ttg gcc atg cag tgg gtg gcg gac aac att gct ggg     624
80 His Asp Gln Arg Leu Ala Met Gln Trp Val Ala Asp Asn Ile Ala Gly
81              195                   200                   205
83 ttt ggc ggc gac ccg agc aag gtg acc ata tac ggc gag tct gcg ggc     672
84 Phe Gly Gly Asp Pro Ser Lys Val Thr Ile Tyr Gly Glu Ser Ala Gly
85      210                   215                   220
87 agc atg tcg acg ttt gtg cac ctt gtg tgg aac gac ggc gac aac acg     720
88 Ser Met Ser Thr Phe Val His Leu Val Trp Asn Asp Gly Asp Asn Thr
89 225                    230                   235                   240
91 tac aac ggc aag ccg ttg ttc cgc gcc gcc atc atg cag tct ggc tgc     768
92 Tyr Asn Gly Lys Pro Leu Phe Arg Ala Ala Ile Met Gln Ser Gly Cys
93                    245                   250                   255
95 atg gtg ccg tct gac ccg gtg gac ggc acg tac ggc acc gag atc tac     816
96 Met Val Pro Ser Asp Pro Val Asp Gly Thr Tyr Gly Thr Glu Ile Tyr
97              260                   265                   270
99 aac cag gtg gtg gcg tct gcc ggg tgt ggc agt gcc agc gac aag ctc     864
100 Asn Gln Val Val Ala Ser Ala Gly Cys Gly Ser Ala Ser Asp Lys Leu
101          275                   280                   285
103 gcg tgc ttg cgc ggc ctt tct cag gac acg ttg tac cag gcc acg agc     912
104 Ala Cys Leu Arg Gly Leu Ser Gln Asp Thr Leu Tyr Gln Ala Thr Ser
105      290                   295                   300
107 gac acg ccc ggc gtg ttg gcg tac ccg tcg ttg cgg ttg tct tat ctc     960
108 Asp Thr Pro Gly Val Leu Ala Tyr Pro Ser Leu Arg Leu Ser Tyr Leu
109 305                    310                   315                   320
111 ccg cgg ccc gac ggc acc ttc atc acc gac gac atg tat gcc ttg gtg     1008
112 Pro Arg Pro Asp Gly Thr Phe Ile Thr Asp Asp Met Tyr Ala Leu Val
113                    325                   330                   335
115 cgg gac ggc aag tac gca cac gtg ccg gtg atc atc ggc gac cag aac     1056
116 Arg Asp Gly Lys Tyr Ala His Val Pro Val Ile Ile Gly Asp Gln Asn
117              340                   345                   350
119 gac gag ggc act ttg ttt ggg ctc tct tct ttg aac gtg acc aca gat     1104
120 Asp Glu Gly Thr Leu Phe Gly Leu Ser Ser Leu Asn Val Thr Thr Asp
121          355                   360                   365
123 gct cag gca cgg gcg tac ttc aag cag tct ttc atc cac gcc agc gat     1152
124 Ala Gln Ala Arg Ala Tyr Phe Lys Gln Ser Phe Ile His Ala Ser Asp
125      370                   375                   380
127 gcg gag atc gac acg ttg atg gcg gcg tac acc agc gac atc acc cag     1200
128 Ala Glu Ile Asp Thr Leu Met Ala Ala Tyr Thr Ser Asp Ile Thr Gln
129 385                    390                   395                   400
```

RAW SEQUENCE LISTING   DATE: 02/25/2005
PATENT APPLICATION: US/09/943,857A   TIME: 14:43:30

Input Set : A:\08919-066001.TXT
Output Set: N:\CRF4\02252005\I943857A.raw

```
131 ggt tct ccg ttc gac acc ggc atc ttc aat gcc atc acc ccg cag ttc      1248
132 Gly Ser Pro Phe Asp Thr Gly Ile Phe Asn Ala Ile Thr Pro Gln Phe
133             405             410             415
135 aaa cgg atc tct gcg ttg ctt ggc gac ctt gcg ttc acg ctt gcg cgt      1296
136 Lys Arg Ile Ser Ala Leu Leu Gly Asp Leu Ala Phe Thr Leu Ala Arg
137             420             425             430
139 cgc tac ttc ctc aac tac tac cag ggc ggc acc aag tac tcg ttc ctc      1344
140 Arg Tyr Phe Leu Asn Tyr Tyr Gln Gly Gly Thr Lys Tyr Ser Phe Leu
141             435             440             445
143 tct aag cag ctt tct ggg ttg ccc gtc ttg ggc acc ttc cac ggc aac      1392
144 Ser Lys Gln Leu Ser Gly Leu Pro Val Leu Gly Thr Phe His Gly Asn
145     450             455             460
147 gac atc atc tgg cag gac tac ttg gtg ggc agc ggc agt gtg atc tac      1440
148 Asp Ile Ile Trp Gln Asp Tyr Leu Val Gly Ser Gly Ser Val Ile Tyr
149 465             470             475             480
151 aac aac gcg ttc att gcg ttt gcc aac gac ctc gac ccg aac aag gcg      1488
152 Asn Asn Ala Phe Ile Ala Phe Ala Asn Asp Leu Asp Pro Asn Lys Ala
153             485             490             495
155 ggc ttg tgg acc aac tgg ccc acg tac acc agc agc tct cag tct ggc      1536
156 Gly Leu Trp Thr Asn Trp Pro Thr Tyr Thr Ser Ser Ser Gln Ser Gly
157             500             505             510
159 aac aac ttg atg cag atc aac ggc ttg ggg ttg tac acc ggc aag gac      1584
160 Asn Asn Leu Met Gln Ile Asn Gly Leu Gly Leu Tyr Thr Gly Lys Asp
161             515             520             525
163 aac ttc cgc ccg gat gcg tac agc gcc ctc ttt tcc aac ccg ccg tct      1632
164 Asn Phe Arg Pro Asp Ala Tyr Ser Ala Leu Phe Ser Asn Pro Pro Ser
165     530             535             540
167 ttc ttt gtg                                                           1641
168 Phe Phe Val
169 545
172 <210> SEQ ID NO: 2
173 <211> LENGTH: 547
174 <212> TYPE: PRT
175 <213> ORGANISM: Candida rugosa
177 <400> SEQUENCE: 2
178 Ser Met Asn Ser Arg Gly Pro Ala Gly Arg Leu Gly Ser Val Pro Thr
179 1               5               10              15
180 Ala Thr Leu Ala Asn Gly Asp Thr Ile Thr Gly Leu Asn Ala Ile Val
181             20              25              30
182 Asn Glu Lys Phe Leu Gly Ile Pro Phe Ala Glu Pro Pro Val Gly Ser
183             35              40              45
184 Leu Arg Phe Lys Pro Pro Val Pro Tyr Ser Ala Ser Leu Asn Gly Gln
185     50              55              60
186 Gln Phe Thr Ser Tyr Gly Pro Ser Cys Met Gln Met Asn Pro Met Gly
187 65              70              75              80
188 Ser Phe Glu Asp Thr Leu Pro Lys Asn Ala Leu Asp Leu Val Leu Gln
189             85              90              95
190 Ser Lys Ile Phe Gln Val Val Leu Pro Asn Asp Glu Asp Cys Leu Thr
191             100             105             110
```

```
192 Ile Asn Val Ile Arg Pro Pro Gly Thr Arg Ala Ser Ala Gly Leu Pro
193         115             120             125
194 Val Met Leu Trp Ile Phe Gly Gly Gly Phe Glu Leu Gly Gly Ser Ser
195         130             135             140
196 Leu Phe Pro Gly Asp Gln Met Val Ala Lys Ser Val Leu Met Gly Lys
197 145             150             155             160
198 Pro Val Ile His Val Ser Met Asn Tyr Arg Val Ala Ser Trp Gly Phe
199             165             170             175
200 Leu Ala Gly Pro Asp Ile Gln Asn Glu Gly Ser Gly Asn Ala Gly Leu
201         180             185             190
202 His Asp Gln Arg Leu Ala Met Gln Trp Val Ala Asp Asn Ile Ala Gly
203         195             200             205
204 Phe Gly Gly Asp Pro Ser Lys Val Thr Ile Tyr Gly Glu Ser Ala Gly
205         210             215             220
206 Ser Met Ser Thr Phe Val His Leu Val Trp Asn Asp Gly Asp Asn Thr
207 225             230             235             240
208 Tyr Asn Gly Lys Pro Leu Phe Arg Ala Ala Ile Met Gln Ser Gly Cys
209             245             250             255
210 Met Val Pro Ser Asp Pro Val Asp Gly Thr Tyr Gly Thr Glu Ile Tyr
211             260             265             270
212 Asn Gln Val Val Ala Ser Ala Gly Cys Gly Ser Ala Ser Asp Lys Leu
213         275             280             285
214 Ala Cys Leu Arg Gly Leu Ser Gln Asp Thr Leu Tyr Gln Ala Thr Ser
215         290             295             300
216 Asp Thr Pro Gly Val Leu Ala Tyr Pro Ser Leu Arg Leu Ser Tyr Leu
217 305             310             315             320
218 Pro Arg Pro Asp Gly Thr Phe Ile Thr Asp Asp Met Tyr Ala Leu Val
219             325             330             335
220 Arg Asp Gly Lys Tyr Ala His Val Pro Val Ile Ile Gly Asp Gln Asn
221         340             345             350
222 Asp Glu Gly Thr Leu Phe Gly Leu Ser Ser Leu Asn Val Thr Thr Asp
223         355             360             365
224 Ala Gln Ala Arg Ala Tyr Phe Lys Gln Ser Phe Ile His Ala Ser Asp
225     370             375             380
226 Ala Glu Ile Asp Thr Leu Met Ala Ala Tyr Thr Ser Asp Ile Thr Gln
227 385             390             395             400
228 Gly Ser Pro Phe Asp Thr Gly Ile Phe Asn Ala Ile Thr Pro Gln Phe
229             405             410             415
230 Lys Arg Ile Ser Ala Leu Leu Gly Asp Leu Ala Phe Thr Leu Ala Arg
231         420             425             430
232 Arg Tyr Phe Leu Asn Tyr Tyr Gln Gly Gly Thr Lys Tyr Ser Phe Leu
233         435             440             445
234 Ser Lys Gln Leu Ser Gly Leu Pro Val Leu Gly Thr Phe His Gly Asn
235     450             455             460
236 Asp Ile Ile Trp Gln Asp Tyr Leu Val Gly Ser Gly Ser Val Ile Tyr
237 465             470             475             480
238 Asn Asn Ala Phe Ile Ala Phe Ala Asn Asp Leu Asp Pro Asn Lys Ala
239             485             490             495
240 Gly Leu Trp Thr Asn Trp Pro Thr Tyr Thr Ser Ser Ser Gln Ser Gly
```

```
241              500              505              510
242 Asn Asn Leu Met Gln Ile Asn Gly Leu Gly Leu Tyr Thr Gly Lys Asp
243         515              520              525
244 Asn Phe Arg Pro Asp Ala Tyr Ser Ala Leu Phe Ser Asn Pro Pro Ser
245      530              535              540
246 Phe Phe Val
247 545
```

```
249 <210> SEQ ID NO: 3
250 <211> LENGTH: 1641
251 <212> TYPE: DNA
252 <213> ORGANISM: Candida rugosa
254 <220> FEATURE:
255 <221> NAME/KEY: CDS
256 <222> LOCATION: (1)...(1641)
258 <400> SEQUENCE: 3
```

```
259 tcg atg aat tca cgt ggc cca gcc ggc cgt ctc gga tcg gta ccc acc    48
260 Ser Met Asn Ser Arg Gly Pro Ala Gly Arg Leu Gly Ser Val Pro Thr
261 1               5                10               15
263 gcc aag ctc gcc aac ggc gac acc atc acc ggt ctc aac gcc atc atc    96
264 Ala Lys Leu Ala Asn Gly Asp Thr Ile Thr Gly Leu Asn Ala Ile Ile
265             20               25               30
267 aac gag gcg ttc ctc ggc att ccc ttt gcc gag ccg ccg gtg ggc aac   144
268 Asn Glu Ala Phe Leu Gly Ile Pro Phe Ala Glu Pro Pro Val Gly Asn
269         35               40               45
271 ctc cgc ttc aag gac cct gtg ccg tac tct ggc tcg ctc aac ggc cag   192
272 Leu Arg Phe Lys Asp Pro Val Pro Tyr Ser Gly Ser Leu Asn Gly Gln
273     50               55               60
275 aag ttc act tct tac ggc ccg tct tgc atg cag cag aac ccc gag ggc   240
276 Lys Phe Thr Ser Tyr Gly Pro Ser Cys Met Gln Gln Asn Pro Glu Gly
277 65               70               75               80
279 acg ttt gaa gag aac ctt ggc aag acg gca ctc gac ttg gtg atg cag   288
280 Thr Phe Glu Glu Asn Leu Gly Lys Thr Ala Leu Asp Leu Val Met Gln
281             85               90               95
283 tcc aag gtg ttc cag gcg gtg ctt ccc cag agt gag gac tgc ctc acc   336
284 Ser Lys Val Phe Gln Ala Val Leu Pro Gln Ser Glu Asp Cys Leu Thr
285             100              105              110
287 atc aac gtg gtg cgg ccg ccg ggc acc aag gcg ggc gcc aac ctc ccg   384
288 Ile Asn Val Val Arg Pro Pro Gly Thr Lys Ala Gly Ala Asn Leu Pro
289         115              120              125
291 gtc atg ctc tgg atc ttt ggc ggt ggg ttt gag atc ggc agc ccc acc   432
292 Val Met Leu Trp Ile Phe Gly Gly Gly Phe Glu Ile Gly Ser Pro Thr
293     130              135              140
295 atc ttc cct ccc gcc cag atg gtc acc aag agt gtg ctc atg ggc aag   480
296 Ile Phe Pro Pro Ala Gln Met Val Thr Lys Ser Val Leu Met Gly Lys
297 145              150              155              160
299 cac atc atc cac gtg gcc gtc aac tac cgt gtt gcc tcg tgg ggg ttc   528
300 His Ile Ile His Val Ala Val Asn Tyr Arg Val Ala Ser Trp Gly Phe
301             165              170              175
303 ttg gct ggt gat gac atc aag gcc gag ggc agc ggg aac gcc ggc ttg   576
```

VERIFICATION SUMMARY   DATE: 02/25/2005
PATENT APPLICATION: US/09/943,857A   TIME: 14:43:31

Input Set : A:\08919-066001.TXT
Output Set: N:\CRF4\02252005\I943857A.raw